Environmental Health Perspectives
Vol. 96, pp. 33–36, 1991

Current Status of the Gene-Tox Program

by Angela E. Auletta,* Mary Brown,† John S. Wassom,† and Michael C. Cimino*

The U.S. Environmental Protection Agency's Gene-Tox Program is a multiphased effort to review and evaluate the existing literature in assay systems available in the field of genetic toxicology. The first phase of the Gene-Tox Program selected assay systems for evaluation, generated expert panel reviews of the data from the scientific literature, and recommended testing protocols for the systems. Phase II established and evaluated the database of chemical genetic toxicity data for its relevance to identifying human health hazards. The ongoing phase III continues reviewing and updating chemical data in selected assay systems. Currently, data exist on over 4000 chemicals in 27 assay systems; two additional assay systems will be included in phase III. The review data are published in the scientific literature and are also publicly available through the National Library of Medicine TOXNET system. The review and analysis components of Gene-Tox comprise 45 published papers, and several others are in preparation. Differences that have been observed between Gene-Tox and National Toxicology Program databases relative to the sensitivity, specificity, accuracy, and predictivity of genetic toxicity data compared to carcinogenesis data are ascribable to differences between the two databases in chemical selection criteria, testing protocols, and chemical class distributions.

Introduction

The U.S. Environmental Protection Agency's Gene-Tox Program is a multiphased effort to review and evaluate the existing literature in genetic toxicology. Phase I of the program was devoted to the selection of assays to be evaluated and the evaluation of literature by work groups of experts in each area. Phase II was devoted to establishing a database of chemicals evaluated by each work group and analyzing that database. Phase III (ongoing efforts) is devoted to the continued review of selected assays and updating of the database, now publicly available through the National Library of Medicine (NLM) TOXNET system. Reports of all three phases are published in *Mutation Research Reviews in Genetic Toxicology* (*1–4*).

Phase I

During phase I of the program, work groups of experts reviewed and evaluated the published literature for 23 selected assays (Table 1) to determine *a*) the validity of a particular system, *b*) the chemicals for which it was best suited, *c*) the proper test protocol, and *c*) the appropriate techniques of data analysis, interpretation, and presentation.

In addition, each work group was asked to *a*) evaluate the assay's ability to discriminate between mutagens and nonmutagens and/or carcinogens and noncarcinogens, *b*) evaluate the system's performance with chemicals of various classes and identify chemicals whose effects were not adequately detected, *c*) formulate generalized protocols and criteria for data evaluation and validation, *d*) identify areas requiring additional research or further development and validation, and *e*) publish an evaluation of the assay in the open literature.

Table 1. Assays evaluated in phase I.

Gene mutation	*Salmonella typhimurium*[a]
	Escherichia coli
	Yeast
	Fungi
	Plants
	Chinese hamster lung cells (V79)[a]
	Chinese hamster ovary cells (CHO)[a]
	Mouse lymphoma L5178Y cells[a]
	Mouse spot test
	Mouse visible specific locus test[a]
Chromosomal effects	Mammalian cytogenetics[a]
	Plant cytogenetics[a]
	Sister chromatid exchange[a]
	Yeast
	Fungi
	Drosophila
	Dominant lethal assay[a]
	Micronucleus assay[a]
	Mouse heritable translocation assay[a]
DNA damage and repair	Repair-proficient and -deficient bacteria
	Unscheduled DNA synthesis[a]
	DNA repair
Oncogenic transformation	Cell strains
	Cell lines
	Viral enhancement
Ancillary assays	Host-mediated assay/body fluid analysis
	Sperm morphology

[a]Assay selected for update.

*U.S. Environmental Protection Agency, Office of Toxic Substances, Health and Environmental Review Division, Washington, DC 20460.
†Human Genome and Toxicology Division, Oak Ridge National Laboratory, Oak Ridge TN 37831.
Address reprint requests to A. E. Auletta, U.S. Environmental Protection Agency (TS-796), Office of Toxic Substances, Health and Environmental Review Division, Washington, DC 20460.

Literature for evaluation was provided to the work groups by the Environmental Mutagen Information Center (EMIC), Oak Ridge National Laboratory, Oak Ridge, Tennessee. EMIC selected only that portion of the available literature which met the following criteria: the article was a primary paper published in a peer-reviewed journal, it dealt with chemical mutagenesis, the agent studied was a pure chemical, the article contained quantitative data, and the article was published in English or a language for which EMIC had easy access to a translation.

Articles that met these criteria were given to the work groups, which then evaluated each article for the following elements: proper use of experimental design; use of positive and negative controls; proper selection of solvents and vehicles; acceptable spontaneous background mutation frequency or rate; use of metabolic activation systems, if necessary; use of appropriate criteria; for positive, negative, or inconclusive results; and provision of dose–response information. This latter criterion was not considered critical if all others were met. In addition, each work group was free to apply other criteria that might be specific to its particular assay. Agents evaluated were designated as positive (+), positive with dose–response data provided (+D), positive with activation only (+*), negative (−), or evaluated but no definitive call possible (T).

At the end of phase I, the work groups had published 37 review articles, 36 concerned with assays in genetic toxicology and 1 describing the establishment of the Gene-Tox carcinogen database (*2*).

Phase II

In addition to the published reports, a database of more than 2000 chemicals had been established at EMIC (*3*). At the outset of the Gene-Tox Program, it was anticipated that this database would be amenable to the type of analysis that would answer a series of fundamental questions about genetic toxicology (Table 2). However, certain characteristics of the database have made such an analysis difficult, if not impossible, to perform.

Chemicals are unevenly distributed across assay systems. For example, of the more than 1000 chemicals in the phase I Salmonella database, approximately 200 had been tested in a cancer bioassay. In comparison, 59 of the approximately 200 chemicals in the mouse lymphoma L5178Y phase I database had been tested in a bioassay.

There is little basis for studies of comparative mutagenesis. In the phase I database, 1559 chemicals, or 59% of the total, had been tested in only one system. Those chemicals that had been tested in more than one system were, for the most part, either direct-acting mutagens or those that are known to metabolize to reactive intermediates by liver enzyme systems. This may have made sensitivity of the various systems appear unnaturally high.

The database is skewed to the positive. With the exception of the Salmonella assay there is a paucity of negative data in the database in general and in the carcinogen database in particular, where only 61 of 506 chemicals evaluated had negative results.

Chemicals tested are unevenly distributed across the 30 classes used in the Gene-Tox classification scheme (Table 3). The most heavily tested classes are class 25, benzene rings; class 30, heterocyclic rings not otherwise classified, unclassified compounds; class 29, alcohols and phenols; class 8, aromatic amines, aliphatic amines, amides, and sulfonamides; and class 2, acyl and aryl halides, halogenated ethers and halohydrins, and saturated and unsaturated alkyl halides. Such a distribution makes an analysis of chemical class specificity of the various assays difficult for all except the Salmonella assay, where a sufficient number of chemicals have been tested in the various classes to permit a determination of system performance according to chemical class.

The phase II analysis resulted in three publications; one dealing with the establishment of the database (*1*), one dealing with the evaluation of mutagenicity assays for the purpose of genetic risk assessment (*3*), and a third dealing with the developmental status of various assays for genetic toxicology testing (*4*).

Table 2. Goals of phase II of the Gene-Tox Program.

Identify genetic and related end points that are of concern to human health
Distinguish those systems that are most ready for extensive use in testing from those that should be regarded as developmental
Determine the sensitivity of each assay to respond to specific classes of chemicals and identify major strengths and weaknesses
Examine the qualitative correlation between mutagenesis and carcinogenesis end points
Devise specialized batteries of bioassays that detect with high probability the various types of genetic and related damage induced by various classes of chemicals
Consider the potential utility of *in vitro* mutagenesis and carcinogenesis bioassays for potency estimation
Identify information gaps and future research needs and establish a mechanism for evaluating the status of test systems on a continuing basis

Table 3. Gene-Tox chemical classification scheme.

1. Acridines, quinacridines, benzimidazoles
2. Acyl and aryl halides, halogenated ethers and halohydrins, saturated and unsaturated alkyl halides
3. Aldehydes, anhydrides
4. Alkyl epoxides, aryl epoxides
5. Alkyl sulfates, sulfoxides, sulfones, sulfonates, organic sulfur compounds not otherwise classified
6. Anthraquinones, quinones
7. Antibiotics, mycotoxins
8. Aromatic amines, aliphatic amines, amides and sulfonamides
9. Aziridines, nitrogen and sulfur mustards
10. Aromatic azo compounds, azoxy compounds, hydrazo compounds, diazoalkanes, nitriles, azides
11. Carbamates, ureas, thioureas, dicarboximides
12. Dioxins, xanthenes, thioxanthenes, phenothiazines
13. Halogens and inorganic derivatives, sulfur and nitrogen oxides
14. Hydrazides, hydrazines, triazenes
15. Hydroxylamines, amine-*N*-oxides
16. Lactones, organic peroxides
17. Mineral fibers
18. Nitroimidazoles, nitrofurans, nitroquinolines, nitroaromatics, nitroalkanes
19. Nitrosamides, nitrosoureas, nitrosoguanidines
20. Nitrosamines
21. Organolead, organomercury, organophosphorus compounds, metals and derivatives, phosphoric acid esters, and phosphamides
22. Polycyclic aromatic hydrocarbons, fluorenones
23. Pyrimidine derivatives, purine derivatives
24. Steroids
25. Benzene ring
26. Amino acids and derivatives
27. Alkaloids
28. Carbohydrates and derivatives
29. Alcohols and phenols
30. Heterocyclic rings not otherwise classified, unclassified compounds

Phase III

As part of the ongoing Gene-Tox effort, certain assays from phase I have been selected for update (Table 1). In addition, two assays not evaluated in phase I, the Chinese hamster ovary (CHO) AS52 assay and the mouse biochemical specific locus assay, will be included in the updated database.

Although the update process has been simplified over that used in phase I, the overall objectives of the program and the basic work group structure remain in place. More than 1500 chemicals have been added to the database since the completion of phase I, bringing to over 4000 the total number of chemicals evaluated. The basic features of the phase III database are the same as those noted above for phase I. There is still a paucity of negative data; the majority of the chemicals evaluated have been tested in only one system and chemical class distribution is essentially unaffected.

The database for the Salmonella assay now totals 2469 chemicals. Of these, 1100 are positive, 880 are negative and 489 are classified as T. Of the 1100 chemicals that are positive, 666 are positive without activation, 416 are positive with activation, and 18 are positive without activation and negative with activation (Table 4).

Of the 2469 chemicals in the Salmonella subset, 328 have associated carcinogenicity data, 268 are classified as carcinogens, and 58 are classified as noncarcinogens. Two hundred ten of the 268 carcinogens are positive in Salmonella; 58 are negative. Of the 58 noncarcinogens, 38 are negative in Salmonella; 20 are positive. Sensitivity is 78%, specificity is 65%, accuracy (concordance) is 76%. Positive predictivity is 91%, negative predictivity is 39%. Zeiger et al. (*5*), reporting on results of the National Toxicology Program (NTP) testing initiative with 114 chemicals, reported 52% sensitivity, 91% specificity, 62% concordance, 90% positive predictivity, and 55% negative predictivity for the Salmonella assay (Table 5).

The Gene-Tox and NTP databases are different in several important aspects. Most notably, chemicals in the NTP were selected according to defined criteria and tested according to standard protocols, whereas chemical selection in Gene-Tox is random, and protocols are varied. In the case of the Salmonella assay, however, the most likely reason for the reported differences in sensitivity, specificity, predictive ability, and concordance of the assay is probably related to chemical class distribution of the agents tested.

The Gene-Tox chemical classification scheme is based on selected organic functional groups, ring systems, biological origins, and/or organic composition. Carcinogens that have been tested in the Salmonella assay are more apt to be classified as halides, epoxides, sulfur compounds, mustards, xanthenes, nitro and nitroso compounds, nitrosamines, metals, polycyclic aromatic hydrocarbons (PAH), steroids, and benzene rings.

Table 4. The Gene-Tox Salmonella database.

Total number evaluated	2469
Positive	1100
Positive without activation	666
Positive with activation	416
Positive without activation, negative with	18
Negative	880
No definite call	489

False positive results are distributed across the data base in the following pattern: alkyl halides, 12; vinyl/allyl compounds 6; halogenated benzenes, benzeneamines and steroids, 5 each; metals and aromatic azides, 4 each; benzene/phenols and ureas/carbamates, 3 each; amides and hydrazines, 2 each; and miscellaneous, 1.

Distribution of the first set of 73 NTP chemicals (*6*) across the Gene-Tox chemical classification scheme shows a relatively high number of agents classified as alkyl halides, allyl and vinyl alkenes, benzeneamines, and aromatic azo compounds; the same classes in which Gene-Tox has found a high proportion of false negative responses. If this distribution holds true for the combined set of 114 chemicals, it could account for the lower sensitivity observed by the NTP and accordingly the differences noted in the other parameters.

Noncarcinogens in the Gene-Tox Salmonella database are found primarily in classes 2 (halides), 8 (aromatic amines), 11

Table 5. Comparison of the Gene-Tox and Salmonella databases.

Database	Sensitivity	Specificity	Accuracy (concordance)	Positive predictivity	Negative predictivity
Gene-Tox	210/268	38/58	248/326	210/230	38/96
	78%	65%	76%	91%	39%
NTP[a]	35/67	43/47	75/114	35/39	43/78
	52%	91%	66%	90%	55%

[a]From Zeiger et al. (*5*).

```
0.        **   ADMINISTRATIVE INFORMATION

          GTN       GENE-TOX Number (Sequential Order)
          DATE      Last Revision Date
          RLEN      Record Length
          UPDT      Update History

1.   ID   **   SUBSTANCE IDENTIFICATION

     NAME      Name of Substance
     RN        CAS Registry Number
     SY        Synonyms
     CCAT      Chemical Classification Category

2.   MSTU **   MUTAGENICITY STUDIES

     GENB      GENE-TOX Evaluation B (Post 1980)

               [Species/Cell Type]
               [Sex]
               [Assay Type]
               [Assay Code]
               [Results]
               [Activation]
               [Dose Response]
               [Reference]
               [Panel Report]

     GENA      GENE-TOX Evaluation A (Pre 1980)

               [Species/Cell Type]
               [Sex]
               [Assay Type]
               [Assay Code]
               [Results]
               [Activation]
               [Dose Response]
               [Reference]
               [Panel Report]
```

FIGURE 1. TOXNET Gene-Tox unit record.

```
GTN      - 14
UPDT     - Complete Update on 11/21/90, 6 Fields
           Added/Edited/Deleted
RLEN     - 1593
NAME     - FORMALDEHYDE
RN       - 50-00-0

GENB
    SPECIES/CELL TYPE    : Chinese hamster ovary (CHO)
                           cells
    ASSAY TYPE           : Gene mutation at the HGPRT
                           locus
    ASSAY CODE           : CHOT
    RESULTS              : No conclusion
    REFERENCE            : EMIC/53976; J TOXICOL
                           ENVIRON HEALTH 12:27-38,
                           1983
    PANEL REPORT         : EMIC/71517; MUTAT RES
                           196:17-36, 1988

GENB
    SPECIES/CELL TYPE    : Mammalian polychromatic
                           erythrocytes, all species
    ASSAY                : Micronucleus test
    ASSAY CODE           : MNTT
    RESULTS              : No conclusion
    REFERENCE            : EMIC/41641; MUTAT RES 90:91-
                           109, 1981
    PANEL REPORT         : EMIC/77345; MUTAT RES
                           239:29-80,1990

GENA
    SPECIES/CELL TYPE    : Neurospora crassa
    ASSAY TYPE           : Reverse mutation
    ASSAY CODE           : NER+
    RESULTS              : Positive
    PANEL REPORT         : EMIC/52327; MUTAT RES
                           133:87-134, 1984
```

FIGURE 2. Gene-Tox unit record for formaldehyde.

(carbamates and ureas), 18 (nitro compounds), 22 (PAH), and 25 (benzene rings). It appears from this analysis that the Salmonella assay can serve as a useful tool for identifying *in vivo* carcinogens, providing attention is paid to the importance of chemical class when interpreting results.

The Gene-Tox database is now publicly available through the National Library of Medicine's TOXNET system. The TOXNET unit record for the Gene-Tox database is shown in Figure 1; a partial record for formaldehyde is shown in Figure 2. Update of the database now that it is public will be primarily the responsibility of the EPA with the assistance of EMIC. The update will continue to make use of the peer-review system although in a slightly modified form. Chemicals to be added to the database will be published in a series of short papers in *Mutation Research Reviews in Genetic Toxicology*; simultaneously with submission of the manuscript, the chemicals evaluated for each assay will be added to the publicly available database. At present, the TOXNET database contains all of those chemicals evaluated in phase I and results of the update for the CHO/HGPRT assay and the micronucleus assay.

The EPA wishes to take this opportunity to express its gratitude to those members of the genetic toxicology community who have given so generously of their time and talent to contribute to the success of the program. We also thank the staff of EMIC for their unfailing support without which this program would not be possible.

REFERENCES

1. Ray, V. A., Kier, L. D., Kannan, K. L., Haas, R. T., Auletta, A. E., Wassom, J. S., Nesnow, S., and Waters, M. D. An approach to identifying specialized batteries of bioassays for specific classes of chemicals: class analysis using mutagenicity and carcinogenicity relationships and phylogenetic concordance and discordance patterns 1. Composition and analysis of the overall data base. Mutat. Res. 185: 197–241 (1987).
2. Nesnow, S., Argus, M., Bergman, H., Chu, K., Frith, C., Helmes, T., McGaughy, R., Ray, V., Slaga, T. J., Tennant, R., and Weisburger, E. Chemical carcinogens. A review and analysis of the literature of selected chemicals and the establishment of the Gene-Tox Carcinogen Data Base. Mutat. Res. 185: 1-195 (1987).
3. Russell, L. B., Aaron, C. S., de Serres, F., Generoso, W. M., Kannan, K. L., Shelby, M., Springer, J., and Voytek, P. Evaluation of mutagenicity assays for purposes of genetic risk assessment. Mutat. Res. 134: 143–157 (1984).
4. Brusick, D., and Auletta, A. Developmental status of bioassays in genetic toxicology. Mutat. Res. 153: 1-10 (1985).
5. Zeiger, E., Haseman, J. K., Shelby, M. D., Margolin, B. H., and Tennant, R. W. Evaluation of four in vitro genetic toxicity tests for predicting rodent carcinogenicity: confirmation of earlier results with 41 additional chemicals. Environ. Mol. Mutagen. 16(18): 1-14 (1990).
6. Tennant, R. W., Margolin, B. H., Shelby, M. D., Zeiger, E., Haseman, J. K., Spalding, J., Caspary, W., Resnick, M., Stasiewicz, S., Anderson, B., and Minor, R. Prediction of chemical carcinogenicity in rodents from in vitro genetic toxicity assays. Science 236: 933–941 (1987).